DOCUMENT RESUME

ED 222 548 TM 820 701

AUTHOR Zimmerman, Irla L.; Woo-Sam, James M.
TITLE Determining the Best WISC-R Short Forms.
PUB DATE 24 Aug 82
NOTE 8p.; Paper presented at the Annual Meeting of the American Psychological Association (90th, Washington, DC, August, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Intelligence Tests; Test Interpretation; *Test Items; Test Reliability; Test Reviews; *Test Selection; *Test Use; Test Validity
IDENTIFIERS *Wechsler Intelligence Scale for Children (Revised)

ABSTRACT

Two kinds of WISC-R short forms, item reduction and subtest reduction, are reviewed in terms of their ability to meet these criteria of adequacy: a significant correlation between the full scale IQ and the short form IQ, a non-significant difference between the full and short form mean IQ, a low percentage of IQ classification changes resulting from the use of the short form in place of the full test, and equal applicability to normal and referred children. The entire spectrum of skills tapped by the WISC-R is assessed by the item reduction short form. The time saving is worth while, and the validity of the short form full scale IQ adequate. Profile analysis procedures should be avoided on the curtailed subtests procedure in view of their poor validity and reliability. Subtest reduction short forms offer a maximum savings in time according to the number of subtests omitted, with quite adequate validity for the longer versions. The best reason for using the subtest rather than item reduction short form is the ease of administration coupled with the ability to complete the entire test if results are at all equivocal. (Author/PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

American Psychological Association
Convention, August 24, 1982

ED222548

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Determining the Best WISC-R Short Forms

Irla Lee Zimmerman, Ph. D.
James M. Woo-Sam, Ph. D.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
I. L. Zimmerman
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Two kinds of WISC-R short forms, item reduction and subtest reduction, are reviewed in terms of their ability to meet these criteria of adequacy: a significant correlation between the Full Scale IQ and the Short Form IQ, a non-significant difference between the Full and Short form mean IQ, a low percentage of IQ classification changes resulting from the use of the short form in place of the full test, and equal applicability to normal and referred children.

Johnny, in special class, hasn't been tested for three years, and retesting is mandatory.

Mary is possibly eligible for a gifted program, but time for evaluation is limited.

Billy has a school phobia and needs to be evaluated before entering psychotherapy.

In a research study, a control group must be compared with an experimental group which will be given the full WISC-R.

These are typical situations facing examiners, each of which may or may not allow for administration of a short form of the WISC-R. The following paper discusses the variety of short forms, what they have to offer, objections to their use, and choices which can be made by examiners in order to answer specific diagnostic or research questions.

Basically, there are two kinds of WISC-R short forms: item reduction, where each subtest is administered, but items on most subtests are limited to every second or third, and subtest reduction, where a specific number of subtests are omitted. Each will be discussed in turn, focusing on the issues of validity, reliability, and application. Factors to be covered are the following: the correlation between full and short form IQ, the difference between full and short form IQ, the percentage of IQ classification changes resulting from the use of the short form, and finally, the applicability of the short form for both normal and referred children.

Item Reduction Short Forms

This method of shortening the entire Wechsler Scale originated with Satz and Mogel (1962), who initially applied it to the WAIS. The idea is to retain the clinical features provided by making use of all the subtests while still reducing the administarion time. Such forms typically utilize 45% of the total test, and testing time is cut in half.

Yudin (1966) adapted the item reduction method to the WISC, and was shortly followed by the Satz group (Satz, Van de Reit, and Mogels, 1967). Differences between short

TM 820 701

forms of this type are based on a) the subtests shortened, b) the specific items to be deleted, and c) the method of prorating to obtain full scale scores. Yudin's procedure is to utilize only the odd items on five of the subtests, only even items on one subtest, and every third item on three subtests. Digit Span and Coding remain unchanged from the original version. The Satz variation is identical to Yudin's except that the correction factors suggested by Yudin are eliminated. Other variations include that of Finch, Ollendick, and Ginn (1973), who also used every third item for three subtests, but began each with #1, and used different correction factors. Silverstein (1968) used a format differing from Yudin's by using odd items on six rather than five subtests and dropping the correction factors. Also, since the Digit Span subtest was not used in the IQ computation in the standardization of the WISC or WISC-R, Silverstein omitted it, as have other researchers since. Hobby (1980) used a split half approach, administering alternate items for all subtests but Coding, an approach also used by Hackett (1979).

Finally, Goh (1979) reported yet another variation. While conforming to the Yudin method of item reduction, he argued that selection of the items should be on a rational basis rather than by arbitrarily selecting odd items, even items, or every third item. In effect, he pleaded the case for local norms, by basing selection on results for a sample of the population to be studied. He reasoned that a short form devised for a specific population will be more accurate than one derived for a general population, and then applied to a specific population.

The validity of the Yudin type short form was estimated for normal children by Silverstein (1982), who based his study on data from the WISC-R standardization sample. As noted above, his formula utilized every other item on six of the subtests, every third on three others, and all of Coding. This formula is comparable for most versions of the item reduction short form. He reported the validity of the short form FIQ as .92, with a standard error of estimate (SE_e) of 5.8.

This encouragingly high correlation is confirmed in a series of research studies covering the full spectrum of bright normal to referred borderline subjects. Specifically, Rasbury, Falgout, and Perry (1978), utilizing 70 bright 7½ year olds, found the validity of the short form FIQ to be .95, SE_e 3.7. Mean IQ differences were insignificant (121 versus 125). However, classification changes were not systematically explored.

Goh (op cit) also utilized a sample of (142) bright youngsters. Both a Yudin type and his own population-specific short forms were calculated, and revealed equivalent validities (.95 and .94). However, mean IQ based on the Yudin format dropped significantly (from a full scale IQ of 110 to 104), while Goh's short form mean IQ was similar to the full scale IQ. When changes in IQ classification were determined, these occured for 67% of the Yudin scores, and only 55% of the Goh scores.

For 100 normal children, Dean (1977) reported (corrected) correlations of .92 if the Yudin scoring were used, or .91 if Silverstein's version were employed. He did not report IQ differences or classification changes. Learning disabled children (N 76) were utilized by Resnick (1979), who reported a validity coefficient of .98. Only 34% of the cases changed classification, but the mean IQ was higher on the short form by three points.

Borderline children were examined by Finch, Kendall, Spirito, Entin, Montgomery, and Schwartz (1979), who reported a correlation of .94 for the Yudin short form (N 100). However, the full scale IQ dropped significantly from 79 to 71, and 50%

of the cases changed intellectual classification. Using a similar sample, Preston (1978) used three separate item reduction short forms: Yudin, Satz, and Finch. Correlations were identical and high (.97), but IQs varied according to the form used. The full IQ was closely matched by the Satz and Finch IQs (78 and 77 respectively), while the Yudin IQ was 7 points lower.

Since all the above studies compared a short form extracted from a standardly administered test, Hackett (1979) explored a separately administered short form with a full administration, using a mentally retarded sample. Although the IQs were identical, the correlation of the extracted short form and the full test was .96, as compared to a correlation of .78 for the separately administered short form and the full test.

Still another issue is this: how valid are the profiles derived from short forms? In other words, how well do item reduction short forms assess the subtests themselves?

For normal children, Dean (op cit) reported that the validity of the nine abbreviated subtests ranged from .46 to .95, with only two above .80 (Yudin form), and from .41 to .90 with four above .80 (Silverstein form). He also compared the internal consistency (split half) of the abbreviated subtests, finding only two to be above .80. Obviously, most subtests were unsatisfactorily estimated in terms of validity and reliability. More encouraging results were seen for bright children: both Rasbury et al (op cit) and Goh (op cit) reported six (Yudin) or seven (Silverstein) of the nine curtailed subtest correlations to be above .80. However, for referred children, high subtest correlations were countered by significant mean differences between the short and long form subtests: Finch et al (op cit) found significant mean differences for eight subtests (usually lower), while Resnick (op cit) also reported significant differences for eight subtests (usually higher).

In essence, the entire spectrum of skills tapped by the WISC-R is assessed by the item reduction short form, a situation which might be particularly desirable for research studies involving a comparison with the full test. Also such an approach might be recommended when intratest scatter would make certain (perhaps undetermined) subtests poor predictors. The time saving is worth while, and the validity of the short form full IQ adequate (.91 or above in the typically extracted short form studies). However, those who wish to indulge in profile analysis should avoid such a procedure on the curtailed subtests, in view of their poor validity and reliability.

Subtest Reduction Short Forms

This method of abbreviating the Wechsler, by omitting subtests, was initially the prototype of short forms. Many proliferated for the WISC, developed more or less by whim. However, soon after the introduction of the WISC-R, a systematic approach to determining the best subtest reduction short forms was presented by Silverstein (1975). A formula that takes subtest unreliability into account was applied to the

average subtest intercorrelations and reliabilities for the 2,200 children in the WISC-R standardization sample. The best combinations of two to five subtests (dyads to pentads), as determined by the extent that each correlated with the full scale, were determined.* Clearly the relationship to the Full IQ increases with the number of subtests used. For example, the best dyad (Vocabulary - Block Design) correlates .88, while for the best pentad (Information - Similarities - Vocabulary - Picture Completion - Block Design), the correlation rises to .93. The relationship between the size of the correlation and prediction of the full scale score is revealed by the standard error of estimate, which, as expected, decreases with the magnitude of the correlation. As calculated by Silverstein, the standard error of estimate for the single best dyad, triad, tetrad, and pentad are 7.1, 6.4, 6.0, and 5.7 IQ points respectively. With errors of estimate of this size, the theoretical percentage of agreement of short form and full scale IQ for subjects falling with Wechsler's seven category system, from very superior to mentally deficient, are 61% for the dyad, 64% for the triad, 67% for the tetrad, and 69% for the pentad. Thus even the best dyads to pentads would misclassify 39 to 31% of the subjects. However, when compared to the item reduction short forms, the best tetrads and pentads appear at least as good, with no greater expenditure of time, and the advantage of a "standard" rather than an altered subtest, as used.

Only one study to date has systematically explored the applicability of the subtest reduction short forms which Silverstein generated to a referred population. Petit (personal communication), using a sample of 210 children certified as perceptually impaired, determined validity scores for each of Silverstein's best combinations. Coefficients were systematically lower than those from the standardization sample, particularly for dyads, which dropped an average of .09. However, pentads were encouragingly close, with less than a .01 difference.

Surprisingly few other studies have explored the Silverstein combinations with the exception of the best "short-short" Vocabulary - Block Design form (r .88). For two large samples of bright children, Killan and Hughes (1978) reported a validity coefficient of .92, while Dirks, Wessels, Quarforth, and Quenon (1980) found a correlation of only .79, with the short form underpredicting the full IQ by (a significant) four points. They attributed the weakness of this combination as due to the erratic inflating effects on scores of exposure to Koh's blocks on the game "Trac Four". For a behavior problem sample, Finch et al (op cit) reported a correlation of .88, and identical mean IQs. However, 44% of the children were misclassified by the short form. (Petit's correlation for this combination was .80).

Preston (1978) explored one of Silverstein's best quartets (I-V-PA-BD) for a borderline population, and found a validity coefficient of .93, or slightly better than Silverstein's .91. While IQs were underestimated by two points, 47% of the sample were within three points of the full IQ. (For the same combination, Petit reported a correlation of .80.)

Most other combinations studied were either arbitrary in selection or drawn from early WISC research. Dirks, et al (op cit) reported various dyads (r .62 to .82),

---

* Sattler (1980) presented similar findings using a slightly different formula. Goh (1979) considers the two sets of results relatively interchangable.

and triads (r .67 to .85) used to identify gifted children with somewhat limited success. On the other hand, Finch, et al (op cit), using a behavior problem sample, reported on a variety of combinations from dyads to pentads. Correlations ranged from .79 to .86 for dyads, with 54 to 57% misclassified, to .80 - .96 for pentads, with only 29 to 35% misclassified. One pentad earlier determined by Finch et al (1973) for the WISC (I-PC-PA-BD-Co), was also checked by Finch et al (1979), who reported a correlation of .96, and a misclassification of only 32%. Cross validated by Erickson and Byrd (1978), correlations for various samples ranged from .91 to .96.

In a well reasoned approach, Kaufman (1976) developed a tetrad which was based on both rational and statistical criteria, using the standardization sample. His combination of Arithmetic - Vocabulary - Picture Arrangement - Block Design systematically taps both verbal and performance scales. The (corrected) validity coefficient if .91, reliability .92, and misclassification a minimal 28%.

In essence, subtest reduction short forms offer a maximum savings in time according to the number of subtests omitted, with quite adequate validity for the longer versions. Perhaps the best reason for using the subtest rather than item reduction short form is the ease of administration coupled with the ability to complete the entire test if results are at all equivocal.

Conclusions

In spite of many warnings as to their limitations, including those implicit in this review, short forms will be used by time pressed examiners. However, knowledgeable selection of an appropriate short form for the task involved can enhance prediction.

References

Petit, M. personal communication, 1981

Gaudet, Renee Comparison of Wechsler short form methodologies as applied to the WISC-R for use with below average children. Dissertation Abstracts International, 1979, 39, 5065B

Hackett, E.F. The reliability of a short form of the WISC-R with educably mentally handicapped subjects. Dissertation Abstracts International, 1979, 40,

BIBLIOGRAPHY

Dean, R.S. The validity and reliability of abbreviated versions of the WISC-R. Journal of Educational and Psychological Measurement, 1977, 37, 1111-1115.

Dirks, Jean, Wessels, Kathleen, Quarforth, Joanne, and Quenon, Betty. Can short-form WISC-R IQ test identify children with high full scale IQs? Psychology in the Schools, 1980, 17, 40-46.

Erikson, Carol L. & Byrd, S.H. Suitability of the California abbreviated WISC for WISC-R subtests with mentally retarded and slow learning children. Psychology in the Schools, 1978, 15, 498-502.

Finch, A.J. Jr., Kendall, P.C., Spirito, A., Entin, A., Montgomery, L.E. & Schwartz, D. J. Short form and factor analytic studies of the WISC-R with behavior problem children. Journal of Abnormal Child Psychology, 1979, 7, 337-344.

Finch, A.J. Jr., Ollendick, T.H., & Ginn, F.W. WISC short forms with mentally retarded children. American Journal of Mental Deficiency, 1973, 78, 144-149.

Goh, D.W. Empirical versus random item selection in the design of intelligence short forms -- the WISC-R example. Applied Psychological Measurement, 1979, 3, 75-82.

Goh, D.W. Note on selection of WISC-R short forms for different uses. Journal of Clinical Psychology, 1979, 35, 319-321.

Haynes, J.P. Validity of a WISC-R short form with male delinquents. Journal of Clinical Psychology, 1982, 38, 406-407.

Karnes, Rances A. & Brown, K.E. A short form of the WISC-R for gifted students. Psychology in the Schools, 1981, 18, 169-173.

Hobby, K.L. WISC-R Split-Half Short Forms. Los Angeles, Western Psychological Services, 1980.

Kaufman, A.S. A four test short form of the WISC-R. Contemporary Educational Psychology, 1976, 1, 108-116.

Kaufman, A.S. Intelligent Testing with the WISC-R. New York, Wiley Interscience, 1979.

Killen, Janice B. & Hughes, L.C. A comparison of short forms of the Binet and WISC-R in the screening of gifted referrals. The Gifted Child Quarterly, 1978, 22, 111-115.

Preston, J. Abbreviated forms of the WISC-R. Psychological Reports, 1978, 42, 883-887.

Rasbury, W.C., Falgout, Janet C., & Perry, N.W. A Yudin-type short form of the WISC-R: Two aspects of validation. Journal of Clinical Psychology, 1978, 34, 120-128.

Resnick, R.J. An abbreviated form of the WISC-R: Is it valid? Psychology in the Schools, 1977, 14, 426-429.

Resnick, R.J. & Entin, A.D. Is an abbreviated form of the WISC valid for Afro-American children? Journal of Consulting and Clinical Psychology, 1971, 36, 97-99.

Ryan, J.J. Clinical utility of a WISC-R short form. Journal of Clinical Psychology, 1981, 37, 389-391.

Sattler, J.M. Assessment of Children's Intelligence and Special Abilities, Second Edition. Boston, Allyn and Bacon, 1982.

Satz, P. & Mogel, S. An abbreviation of the WAIS for clinical use. Journal of Clinical Psychology, 1962, 18, 77-79.

Satz, P., Van De Riet, H., & Mogel, S. An abbreviation of the WISC for clinical use. Journal of Consulting Psychology, 1967, 31, 108.

Silverstein, A.B. A short-short form of the WISC-R for screening purposes. Psychological Reports, 1974, 35, 817-818.

Silverstein, A.B. Validity of a new approach to the design of the WAIS, WISC, and WPPSI short forms. Journal of Consulting and Clinical Psychology, 1968, 32, 478-479.

Silverstein, A.B. Validity of Satz-Mogel-Yudin-type short forms. Journal of Consulting and Clinical Psychology, 1982, 54, 20-21.

Silverstein, A.B. Validity of WISC-R short forms. Journal of Clinical Psychology, 1975, 31, 696-697.

Yudin, L.W. An abbreviated form of the WISC for use with emotionally disturbed children. Journal of Consulting Psychology, 1966, 30, 272-275.